



# Étude expérimentale d'une nouvelle mesure de qualité d'image fondée sur les propriétés de l'analyse en ondelettes

Stéfane Paris

## ► To cite this version:

Stéfane Paris. Étude expérimentale d'une nouvelle mesure de qualité d'image fondée sur les propriétés de l'analyse en ondelettes. Ateliers de Traitement et Analyse de l'Information - Méthodes et Applications, 2011, Hammamet, Tunisie. hal-01223730

**HAL Id: hal-01223730**

**<https://hal.science/hal-01223730>**

Submitted on 4 Nov 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Étude expérimentale d'une nouvelle mesure de qualité d'image fondée sur les propriétés de l'analyse en ondelettes

Sté fane Paris

Département informatique, UFR MIM,, Qgar, LORIA,,  
Université de Lorraine, France, *email*: Stefane.Parisloria.fr

## Abstract

Cet article présente une étude expérimentale d'une mesure objective de qualité d'images connaissant l'image de référence et reposant sur l'analyse en ondelettes. La décomposition en ondelettes offre une représentation multiéchelle orientée qui est prise en compte par l'utilisation d'une généralisation de la distance euclidienne classique. Une étude statistique est faite à l'aide d'un ensemble de bases d'images de référence dans le domaine.

**Mots clés** Mesure de qualité d'images, analyse en ondelettes, distance euclidienne généralisée.

## Abstract

This paper presents an experimental study of an objective image quality assessment with full reference (FR-IQA) based on wavelet decomposition. The wavelet analyses provide a multiscale oriented representation which is observed through a generalized Euclidean distance to achieve the proposed FR-IQA. Statistical evaluations are performed on several image databases.

**Key words** Image quality assessment, wavelet analysis, wavelet-based distance.

## 1 Introduction

Les transformations multiéchelles sont des modèles assez fidèles du système visuel humain (SVH) pour un coût algorithmique relativement raisonnable ; parmi celles-ci on trouve les filtres de Gabor et les ondelettes. Aussi, la qualité d'images connaissant l'image de référence (QI-IR) peut être faite à l'aide d'une analyse en ondelettes. partir de la différence des analyses en ondelettes de l'image de référence et de l'image déformée, la moyenne des énergies pondérées des sous-bandes fournit une mesure acceptable dans une première approche. La pondération des sous-bandes a été étudiée expérimentalement par Watson *et al.* [7] et consiste à ignorer les hautes fréquences pour éviter les perturbations dues aux bruits et à pondérer les autres sous-bandes de fa on à correspondre à la perception humaine. Cependant, deux approches majeures dans le domaine,

VIF<sup>1</sup> et SSIM, proposent une mesure plus fine de la QI-IR. En effet, les auteurs de ces travaux ont montré que l’œil humain est plus sensible aux détériorations lorsque celles-ci apparaissent dans des zones de l’image texturées. Ainsi, SSIM évalue la qualité d’une image en observant ses déformations *locales* de structure, de contraste et de luminosité à l’aide de moyennes, variances et corrélations. Utilisant une modélisation par champs aléatoires gaussiens du SVH, VIF décrit deux types de bruit : le bruit additif et l’atténuation par floutage. La qualité est alors le rapport entre les informations mutuelles (IM) associées à la perception humaine de l’image de référence d’une part et de l’image déformée d’autre part. Le cumul des IM de chaque sous-bande fournit une mesure globale de la qualité. Ces deux algorithmes de QI-IR définissent leurs mesures localement et donc sous-tendent l’utilisation d’un fenêtrage qui est déjà implicitement défini dans la décomposition dyadique en ondelettes.

Cet article propose une approche de la QI-IR fondée sur les propriétés quelque peu délaissées de l’analyse en ondelettes et ce grâce à l’introduction de la distance euclidienne généralisée telle que l’a défini Z. Wang *et al.* [2]. Cette distance entre les image de référence et déformée a pour unité de mesure le pixel caractérisé par un vecteur de coefficients d’ondelettes provenant des différentes sous-bandes. Par la suite l’article est organisé comme suit. La section 2 présente la caractérisation des pixels et ses avantages ainsi que la distance euclidienne généralisée associée. La section 3 décrit plusieurs expérimentations utilisant différents banc de filtres pour l’analyse en ondelettes et différents noyaux pour la distance euclidienne généralisée. Un ensemble de statistiques est évalué et comparé aux statistiques fournies par plusieurs principes de QI-IR à l’aide de sept bases d’images dédiées à cette tâche [3, 8].

## 2 Distance euclidienne généralisée – WED

L’analyse en ondelettes étant dyadique, un coefficient d’une sous-bande à faible résolution est commun aux quatre coefficients de la sous-bande de résolution directement supérieures. Ainsi, la corrélation inter pixels est automatiquement décrite par une analyse en ondelettes (cf. Fig. 1). Chaque image est décrite par un ensemble de descripteurs qui chacun informe sur les coefficients d’ondelettes d’un pixel. La distance euclidienne généralisée (WED) est alors utilisée pour renforcer la similarité entre pixels ayant des coefficients communs :  $d_{WED}^2(x, y) = \sum_{i,j=1}^n g_{i,j}(x_i - y_i)(x_j - y_j)$ . Le noyau  $g_{i,j}$  est défini positif, continu [2]. On utilisera par la suite la gaussienne comme noyau :  $g_{i,j} = \exp(-(i - j)^2 / (2\sigma^2)) / \sqrt{2\pi\sigma^2}$  dans une version simplifiée où  $\sigma = 1$  et la constante de normalisation est ignorée ; ces simplifications n’apportant aucune modification des propriétés requises pour utiliser la distance. En particulier, une décroissance rapide de la fonction noyau est assurée en fixant  $\sigma$  à 1 :  $g_{i,j} \approx \exp(-(i - j)^2 / 2)$ . Ceci afin de palier à l’incapacité de la distance euclidienne classique d’être en adéquation d’avec le SVH. L’exemple de Z. Wang *et al.* [5] où les images sont mesurées à égale distance, au sens des moindres carrés, de l’image de référence ne sont pas per ues avec le même taux de déformation. Pour de plus amples détails sur les propriétés de WED dans [4].

<sup>1</sup>Par soucis d’espace limité, les références aux bases d’images et aux algorithmes de QI-IR se feront indirectement via les les états de l’art [3, 8].

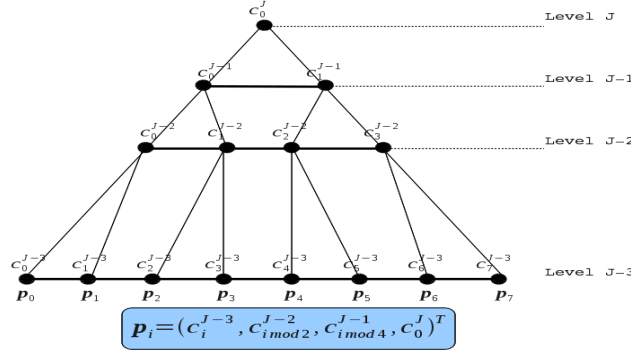


Figure 1: Une représentation arborescente de l'analyse en ondelettes à trois niveaux ( $J = 3$ ) d'un signal mono-dimensionnel montrant les coefficients de sous-bandes que partagent les échantillons: le coefficient  $C_1^{J-2}$  est commun aux pixels  $p_2$  et  $p_3$ ; le pixel  $p_2$  est décrit par le descripteur  $(C_0^J, C_0^{J-1}, C_0^J, C_1^{J-2}, C_2^{J-3})^T$  et  $p_3$  par  $(C_0^J, C_0^{J-1}, C_0^J, C_1^{J-2}, C_3^{J-3})^T$ .

### 3 Étude expérimentale

La distance WED a été testée sur plusieurs bases d'images dédiées à la QI-IR : LIVE, CSIQ, IVC, Toyama, A57, TID et WIQ. Guidé par les états de l'art de W. Lin *et al.* et de F. Zhang *et al.* [3, 8], WED est évaluée et comparée aux algorithmes de QI-IR connus : SSIM, VIF, MSVD, IFC, VSNR et PSNR. Le tableau 1 montre les coefficients de corrélation de Pearson ( $C_P$ ) et d'ordre ( $C_S$ ) pour ces différentes mesures de QI-IR. VIF apparaît comme étant le meilleur avec des coefficients significatifs sauf pour la base A57. SSIM est souvent classé deuxième avec des coefficients proches et une complexité algorithmique nettement moindre. La qualité mesurée par WED est faite en utilisant le banc CDF9-7 connu pour offrir une excellente qualité perceptuelle des sous-bandes. Les coefficients de corrélation sont rangés dans la table 2 où plusieurs configurations ont été testées :

**Ligne 1**  $J=3$  et l'ensemble de la décomposition est prise en compte. C'est la configuration de base ;

**Ligne 2**  $J=3$  et les hautes fréquences sont ignorées comme le préconisent D; Watson *et al.* in [7] ;

**Ligne 3**  $J=3$ , les hautes fréquences sont ignorées et la pondération de Watson *et al.* est utilisée [7] ;

**Ligne 4**  $J=3$  et un prétraitement de sous-échantillonnage similaire à celui de SSIM est utilisé [6] ;

**Ligne 5**  $J=5$  et la configuration de base est testée ;

**Ligne 6**  $J=5$  et le sous-échantillonnage est appliqué.

Globalement, il apparaît que les meilleurs coefficients de corrélation sont obtenus avec la configuration  $J=3 - \mathbf{HF}^2$ . Le tableau 3 compare les coefficients  $C_P$  et  $C_S$  de cette configuration avec ceux de SSIM et de VIF. WED est plus faible pour les bases IVC, Toyama et WIQ. En revanche, ses coefficients sont les meilleurs pour la base CSIQ et ils sont en seconde position – proches des coefficients en première position – pour les autres bases. D'autres tests ont été effectués

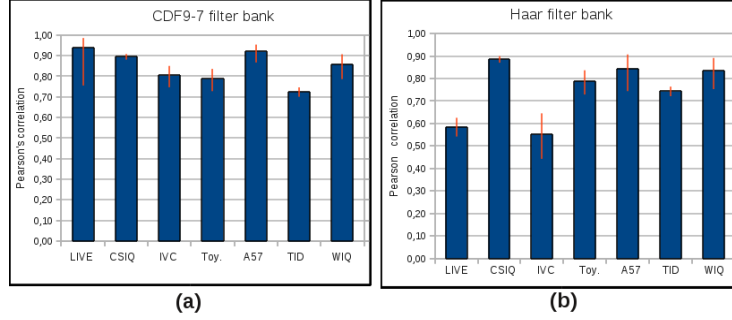


Figure 2: Les coefficients de corrélation de Pearson des bases d'images et leurs intervalles d'acceptation (seuil de rejet à 5).

avec les bancs Haar, DB8 et Coiflet. Alors que les coefficients des bancs DB8 et Coiflet sont nettement inférieurs à ceux du banc CDF9-7, ceux du banc Haar y sont comparables. La figure (2) montre la valeur de  $C_P$  et son intervalle d'acceptation (IA) pour chacune des bases. Le graphique de gauche correspond à l'utilisation du banc CDF9-7 et celui de droite au banc Haar avec, pour les deux bancs, la configuration  $J = 3$  -HF. Les valeurs sont très proches avec un léger affaiblissement pour A57 et WIQ et une perte significative pour LIVE et IVC. De surcroît, les IA sont plus resserrés avec le banc de Haar qu'avec le banc CDF9-7.

Le concept de structure multiéchelle introduit par Koenderink and van Doorn [1], a également été testé comme amélioration possible de WED :  $g_{i,j} = \exp(-((i-j) \div 3)^2/2)$ . Le noyau de WED n'est plus décroissant mais périodique avec des amplitudes égales pour des coefficients de même orientation mais à différentes échelles. Ainsi, cette nouvelle définition de WED est orientée suivant les directions  $0$ ,  $\pi/4$  et  $\pi/2$ . Comme le montre le tableau 4, cette propriété améliore légèrement les performances pour la configuration Haar excepté pour les bases A57, CSIQ et WIQ.

Les cartes de déformations montrées en figure 3 sont obtenues avec la configuration **CDF9-7 J=3 -HF**. Le premier couple montre que la couleur n'est pas prise en compte car le cœur de chaque carré n'est pas considéré comme déformé. (Seule la texture est perue) Avec le couple 5 on constate que la localisation est limitée en précision par l'utilisation de la décomposition en ondelettes. Mais, globalement, il apparaît que l'intensité des pixels de la carte de déformation (image  $i.2$ ) correspond relativement à la perception de la déformation que l'on peut avoir.

Rappelons que WED a pour unité de mesure le pixel et de fait la localisation spatiale de la déformation est estimée en même temps que la mesure globale de qualité est calculée. Un post-traitement utilisant l'amplitude de la déformation en chaque pixel devrait être facilement défini pour être incorporé dans une tâche telle que la compression ou la diffusion sur réseau d'images où la perception humaine est prépondérante.

<sup>2</sup>Le niveau de décomposition maximal est fixé à 3 et les hautes fréquences sont ignorées.

## 4 Conclusion

Cet article vient de montrer l'efficacité algorithmique de WED qui opère simultanément deux opérations. Elle fournit une mesure globale de la qualité d'une image avec des performances correctes voire bonnes suivantes les bases d'images tests observée. Et elle localise les déformations puisqu'elle est fondée sur le concept de distance euclidienne généralisée. Il est à noter que l'incorporation de l'information couleur est à tester dans un objectif d'amélioration des performances. Pa ailleurs, il semblerait que l'amplitude de chaque pixel de la carte de déformation devrait informer sur les zones de forte déformation perceptuelle.

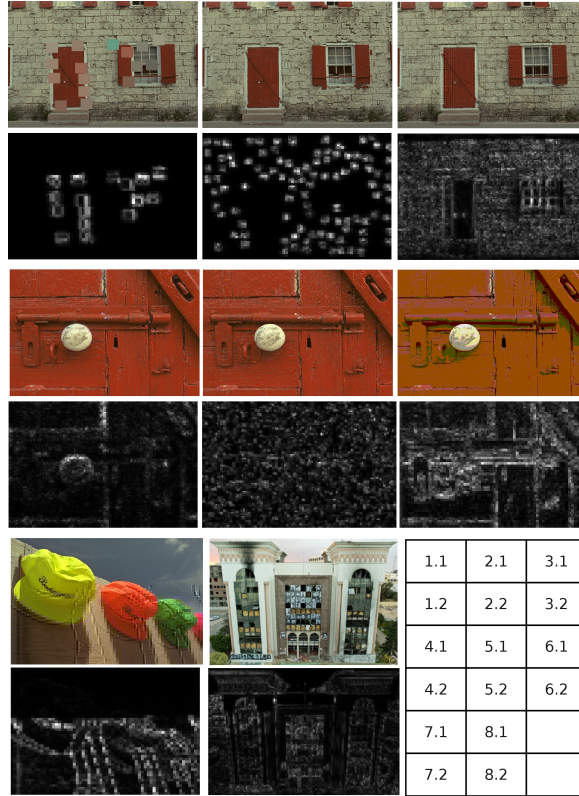


Figure 3: Localisation de la distorsion : chaque couple  $i = (i.1; i.2)$  est composé d'une image déformée ( $i.1$ ) et de sa carte de déformation ( $i.2$ ).

## References

- [1] J.J. Koenderink and A.J. van Doorn. Image structure. In *DAGM-Symposium*, pages 3–35, 1997.
- [2] Wang L., Zhang Y., and Feng J. On the euclidean distance of images. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 27(8):1334–1339, August 2005.

- [3] W. Lin and C. C. Jay Kuo. Perceptual visual quality metrics: A survey. *J. Vis. Commun. Image Represent.*, 22:297–312, May 2011.
- [4] S. Paris. Wavelet-based euclidean distance for image quality assessment. In *Int. Symp. ELMAR*, September 2010.
- [5] Z. Wang and A. C. Bovik. Mean squared error : love it or leave it ? - a new look at signal fidelity measures. *IEEE Signal Processing Magazine*, 26(1):98–117, January 2009. <http://live.ece.uwaterloo.ca/research/quality/>.
- [6] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment : From error visibility to structural similarity. *IEEE Transaction on Image Processing*, 13(4):600–612, April 2004.
- [7] A. B. Watson, G. Y. Yang, J. A. Solomon, and J. Villasenor. Visibility of wavelet quantization noise. *IEEE Transaction on Image Processing*, 6(8):1164–1174, August 1997.
- [8] Fan Zhang, Lin Ma, Songnan Li, and King Ngi Ngan. Practical image quality metric applied to image coding. *Multimedia, IEEE Transactions on*, 13(4):615–624, aug. 2011.

Table 1: Les coefficients de corrélation (avec l’aimable autorisation de W. Lin [3]). En police **grasse** sont indiqués les meilleurs coefficients.

		LIVE	CSIQ	IVC	Toyama	A57	TID	WIQ
$C_P$	PSNR	0.8724	0.7677	0.7196	0.6329	0.6273	0.5677	0.7601
	SSIM	0.9147	0.8563	0.9018	0.8872	0.8033	0.7931	<b>0.7876</b>
	MSVD	0.9029	0.7776	0.7725	0.8149	0.7099	0.6423	0.7433
	VIF	<b>0.9567</b>	<b>0.9242</b>	<b>0.9028</b>	<b>0.9138</b>	0.6139	<b>0.8059</b>	0.7559
	IFC	0.9227	0.8459	0.9056	0.8403	0.4788	0.7249	0.7431
	VSNR	0.9055	0.7978	0.8025	0.8704	<b>0.9210</b>	0.6790	0.7623
		LIVE	CSIQ	IVC	Toyama	A57	TID	WIQ
$C_S$	PSNR	0.8766	0.7941	0.6885	0.6133	0.6189	0.5773	0.6257
	SSIM	0.8991	0.8656	<b>0.9017</b>	0.8795	0.9103	<b>0.7967</b>	<b>0.7261</b>
	MSVD	0.8801	0.7871	0.7734	0.8146	0.6485	0.6520	0.6362
	VIF	<b>0.9577</b>	<b>0.9195</b>	0.8964	<b>0.9077</b>	0.6223	0.7457	0.6918
	IFC	0.9239	0.7722	0.8793	0.8355	0.3987	0.6644	0.7113
	VSNR	0.8975	0.8089	0.7993	0.8609	<b>0.9355</b>	0.7010	0.6558

Table 2: Les coefficients de Spearman pour WED ; **J** : niveau maximal de décomposition ; **-HF** : suppression de Hautes Fréquences [7] ; **+W** : suppression des HF et pondération de Watson [7] ; **+S** : sous-échantillonnage similaire à celui de SSIM. En police **grasse** sont indiqués les meilleurs coefficients.

		LIVE	CSIQ	IVC	Toyama	A57	TID	WIQ
$C_P$	<b>J=3</b>	0.7989	0.6841	0.7623	0.6869	0.8598	0.7783	0.7296
	<b>J=3 -HF</b>	0.9507	<b>0.9934</b>	0.7934	<b>0.8141</b>	0.9104	0.7814	0.7347
	<b>J=3 +W</b>	0.9108	0.6602	<b>0.7984</b>	0.7316	0.8270	0.7760	0.7357
	<b>J=3 +S</b>	0.9232	0.8249	0.7653	0.7901	<b>0.9155</b>	<b>0.7897</b>	<b>0.8090</b>
	<b>J=5</b>	<b>0.9514</b>	0.7684	0.7711	0.6883	0.7824	0.5716	0.7540
	<b>J=5 +S</b>	0.9511	0.7787	0.7783	0.7046	0.7835	0.6812	0.7537
		LIVE	CSIQ	IVC	Toyama	A57	TID	WIQ
$C_S$	<b>J=3</b>	0.9306	0.7649	0.7718	0.7845	0.8161	0.7625	0.6638
	<b>J=3 -HF</b>	<b>0.9351</b>	<b>1</b>	0.8058	0.8188	0.8896	0.7637	0.6705
	<b>J=3 +W</b>	0.9299	0.7757	<b>0.8116</b>	0.8112	0.8203	0.7610	0.6806
	<b>J=3 +S</b>	0.8928	0.8017	0.7607	<b>0.8464</b>	<b>0.9167</b>	<b>0.7725</b>	0.7804
	<b>J=5</b>	0.9343	0.8014	0.5256	0.8184	0.7785	0.6755	0.7837
	<b>J=5 +S</b>	0.9340	0.8018	0.5277	0.83010	0.7832	0.6657	<b>0.7879</b>

Table 3: Comparaison des coefficients de corrélation entre SSIM, VIF et WED\_CDF9-7\_J3-HF.

$C_P$	LIVE	VIF< <b>WED</b> <SSIM	$C_S$	LIVE	VIF< <b>WED</b> <SSIM
	CSIQ	<b>WED</b> <VIF<SSIM		CSIQ	<b>WED</b> <VIF<SSIM
	IVC	VIF<SSIM< <b>WED</b>		IVC	SSIM<VIF< <b>WED</b>
	Toyama	VIF<SSIM< <b>WED</b>		Toyama	VIF<SSIM< <b>WED</b>
	A57	<b>WED</b> <SSIM<VIF		A57	SSIM< <b>WED</b> <VIF
	TID	VIF<SSIM< <b>WED</b>		TID	SSIM< <b>WED</b> <VIF
	WIQ	SSIM<VIF< <b>WED</b>		WIQ	SSIM<VIF< <b>WED</b>



Table 4: Coefficients  $C_S$  de CDF9-7\_WED et Haar\_WED quand la structure multiéchelle est prise en compte.

	<b>LIVE</b>	<b>CSIQ</b>	<b>IVC</b>	<b>Toyama</b>	<b>A57</b>	<b>TID</b>	<b>WIQ</b>
<b>CDF9-7</b>	0.9322	0.7738	0.8002	0.8109	0.8847	0.7588	0.5633
<b>Haar</b>	0.9376	0.7854	0.7966	0.7829	0.7563	0.7761	0.7403